

ON FURTHER DEVELOPMENT OF THE CONCEPT OF SOCIAL VALIDITY

JACK W. FINNEY

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

"WHY DON'T YOU ASK THEM?"

After considerable deliberation about which of several behavioral checklists should be used to determine whether parents reported that their children had behavior problems, I was unsatisfied with my choices. Each checklist gave a behavior-problem score, but none of the checklists told me whether parents thought their children's behavior was a problem. Assuming that there would be some psychometric rationale for choosing one from many behavioral checklists, I consulted with a statistician with psychometric expertise. We reviewed checklists and their reliability and validity data, and clarified the intended use of the checklist. The consultant listened and then made a simple, straightforward suggestion. "If you want to know whether parents think their children have behavior problems, why don't you ask them?" Consequently, I turned my attention from psychometrically sound but assumption-ridden checklists to a simple, direct question to be asked of parents.

Wolf (1978) provided a similar prescription for behavior analysts, recommending they ask their consumers: How satisfied are you with the services, the people who provide them, and the outcomes that have resulted? Wolf provided several examples of using consumer evaluations to modify and refine the programs and procedures of Achievement Place and directed other applied behavior analysts to focus on subjective consumer evaluations to promote optimal development of other behavior-change programs.

Wolf's (1978) paper on social validity changed most behavior analysts' views of how to judge the importance of research findings. Decisions about the importance of a behavior-change technique and/or its outcomes had previously been made by the researchers, journal editors, and reviewers, and by readers who later cited or used the technique. Now other consumers were to be involved in these judgments. Editors and reviewers began to ask for social validity measures as support for objective measures. In fact, the reviewers of one of our manuscripts suggested that the findings were interesting but could only be published if social validity data were obtained. (We obtained the data, and the article was published.) The notion that consumers can provide useful information about the importance of data is becoming ensconced in the methods of applied behavioral research.

"ASK THEM?"

Schwartz and Baer (1991) have provided further refinements to the concept of social validity, with special attention to expanding and classifying the "them" in my earlier question. Direct and indirect consumers of psychological services are familiar groups already included in many social validity assessments. However, members of the immediate and extended communities are significant additions to important populations of consumers. By focusing on a range of consumer groups, Schwartz and Baer provide another dimension along which behavior-change programs can be evaluated socially (see Kazdin, 1977, for some additional dimensions).

If Schwartz and Baer's (1991) suggestions are followed, behavioral researchers will have the task of developing reliable and meaningful placement of consumers into consumer classes. Direct consumers should be easily identified, and perhaps the same is true of indirect consumers—the "brokers" of

Supported in part by a grant (RO1 MH43883) from the National Institute of Mental Health. Reprints may be obtained from Jack W. Finney, Child Study Center, Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0436 (Bitnet address: FINNEY@VTVM1).

behavioral services for others. Members of immediate and extended communities may be more difficult to identify, and, for some types of target problems, one group of consumers may be categorized as one class of consumers in one study but categorized differently by other investigators. For example, the social validity of a program designed to enhance the quality of physician visits for mothers seeking well-child care (Finney et al., 1990) could be evaluated by mothers as direct consumers, by health-care administrators as indirect consumers, by a sample of all users of health-care services at a facility as members of an immediate community, and by health-care administrators of other facilities in the city as members of an extended community. On the other hand, the social validity of a program designed to provide psychological services and thereby reduce excessive health-care use by members of a health maintenance organization (Finney, Riley, & Cataldo, in press) could be evaluated by health-care administrators as direct consumers and by health-care administrators of other facilities as members of an immediate community. Justification for the choices and classifications of consumers will be necessary for each study. Sampling multiple groups of direct, indirect, immediate, and extended consumers within a study may facilitate the determination of guidelines for selecting important consumers and developing reliable and meaningful consumer classes.

Winett, Moore, and Anderson (1991) proposed a broader viewpoint of consumers and social validity by directing attention to epidemiological data for validating the importance of problems, interventions, and program outcomes. Verifiable importance places an emphasis on knowing the incidence and prevalence of disorders, with an implicit assumption that disorders with high incidence or prevalence merit societal resources. Thus, consumers could be defined as citizens of a nation, a state, or a community, and verifiable importance could be related to characteristics and needs of the specific population.

For many health issues, the epidemiological approach leads appropriately to targeting those disorders for prevention or treatment that threaten the

most. This approach, however, should not preclude attention to low-incidence disorders with high severity. Low-incidence disorders merit societal resources when the associated dysfunction is severe and require more resources when prevention or early treatment is not provided. Verifiable importance therefore provides a definition of consumers that is population based, but the size of the population may range from large (e.g., the nation) to small (e.g., elementary students in special education classes). By integrating epidemiological data as beginning and end points for behavior-change programs, Winett et al. (1991) provide directions for selecting targets, interventions, and populations, and for obtaining new indexes of efficacy that relate to optimal outcomes. Although the authors use health-related examples to illustrate this broader definition of social validity, the epidemiological approach has equal applicability to psychological disorders (Costello, 1989), which have recently received greater attention from epidemiology researchers.

SHOULD SOCIAL VALIDITY ASSESSMENT BE REQUIRED?

Behavioral journals might follow Schwartz and Baer's (1991) advice to establish criteria for conducting and reporting social validity assessments as a requirement for publication. In fact, there is anecdotal evidence that social validity criteria are at least sometimes used currently for manuscript review, with the implication that supportive social validity data are needed. Further development of these criteria, and encouragement for assessment of social invalidity, should include procedures to minimize a publication bias favoring studies with high scores on social validity measures. Journals most often publish studies that result in success, with failures only occasionally surviving editorial reviews.

What will happen to studies with low social validity data or high social invalidity data during the review process? Failures, when obtained with sufficient attention to avoiding Type II errors, can be highly instructive. Similarly, we can learn from programs with poor consumer validity. Behavioral journals should provide clear guidelines to authors

and reviewers for judging the importance of studies that investigate problems, interventions, and outcomes with low social validity or high social invalidity. These latter studies may investigate important but currently underappreciated problems or behavior-change strategies that will lead to new research programs on topics that later become valid and verifiably important.

Innovations in behavioral research may often appear in the literature as case reports, brief reports, and observational studies. Should social validity assessments be required for these early reports? McKinlay (1981) reviewed a number of medical innovations that seemed quite promising, but later were found to be ineffective, costly, or socially unacceptable when randomized controlled trials were conducted. It may therefore be critical to determine social validity in initial, innovative studies, with care taken not to abandon a promising intervention due to poor consumer reactions. Perhaps not all studies should be required to provide social validity assessments, but not knowing consumers' reactions can pose obstacles to further development of an intervention and lead to ineffective programs (Winnett et al., 1991; Wolf & Ramp, 1991). As society strives to provide equitable access to physical and mental health interventions through public funding, rigorous criteria of effectiveness, cost effectiveness, and social acceptance for medical (and psychological) innovations will be implemented (McKinlay, 1981). Behavioral research has valued the fulfillment of these three criteria. Future research can provide conceptual and pragmatic knowledge related to the implementation of these criteria, which could be useful to many other fields of study.

ARE NEW METHODS NEEDED?

If behavioral journals require social validity assessment, methodological refinements and advances will be necessary. Determining the reliability and validity of measures of social validity should receive greater attention (Schwartz & Baer, 1991), as should the exploration of new measurement approaches. We may not, however, need to abandon subjective self-report measures. Carefully constructed self-re-

port measures of health, well-being, illness, pain, and other difficult-to-observe behavioral states have been developed (e.g., Diaz et al., 1986; Eisen, Ware, Donald, & Brook, 1979). Some types of social validity may best be measured by self-report (see Hawkins, 1991, for some examples). A focus on the assessment context and questions resulting in accurate self-reports could advance this measurement technology for research and clinical practice.

Procedural and operational guidelines have been delineated to refine old and provide new approaches to social validity assessment (Fawcett, 1991). Hawkins (1991) has suggested that social validity should be based on a functional approach to determine the predictive or habilitative validity of self-report or more objective consumer measures. Statistical approaches have been devised to assess meaningful change for psychotherapy (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991), and the extension and refinement of these techniques for other preventive and therapeutic endeavors should provide additional operations for establishing social validity.

Winnett et al. (1991) suggest a broader conceptualization of social validity for systematic use during formative research (and thereafter) to (a) provide refinement of program features, (b) assist in targeting populations using social marketing strategies, and (c) produce better outcomes. Such use of epidemiological and social marketing approaches as part of social validity assessment warrants attention from behavioral researchers (Baer & Schwartz, 1991). Schwartz (1991) suggests that research in business and consumer education can provide a conceptual framework for further development of social validity from consumers of behavioral programs. From these myriad suggestions for further development of social validity, one can envision research programs focused on social validity and its development. Simultaneously, investigators can apply and investigate some or many of these social validity suggestions in the context of validating programs and outcomes. Both established and new methods and measures are likely to turn up in future social validation efforts.

AN INTERACTIVE PROCESS

Social validation for behavioral research should be an interactive process. A researcher implements and evaluates an intervention for a problem; a consumer (or several classes of consumers) provides an evaluation of the intervention and its outcomes; the researcher changes and improves (but not always) the intervention, which is then reimplemented and subsequently reevaluated by a more experienced consumer; and so on. Behaviors of the researcher and the consumer are changed by this interactive process. To enable informed judgments about behavioral programs, consumers become educated about problems and their importance and about interventions and their outcomes. Researchers are educated by the consumers of their products regarding the good and bad features and the desirable and undesirable outcomes. Model intervention programs developed and evaluated from a dynamic, interactive approach (e.g., Winett et al., 1991; Wolf, 1978; Wolf & Ramp, 1991) provide evidence of the benefits of mutual influence of researcher and consumer. Too often social validity has been a static, one-time measure, obtained only from direct consumers. Additional tests of the interactive model of program development are needed to better define and develop the process now called social validity.

"KEEP ASKING THEM"

Wolf (1978) argued for the importance of subjective measures from consumers to advance the acceptability and effectiveness of behavioral programs. The articles in this issue of the *Journal of Applied Behavior Analysis* present refinements of and new ideas about social validity and its measurement. Inclusion of varied classes of consumers can yield meaningful evaluations. Self-reports of consumers' opinions about behavioral programs and assessments of objective and functional indicators of consumers' use of behavioral services will further the goal of effective, acceptable research and practice. Using consumer information interactively and considering epidemiological data should facilitate the choice of targets and an evaluation of the attainment of behavior-change goals.

Behavioral researchers are challenged by many problems to be solved, interventions to be developed, and outcomes to be assessed. They can and should be assisted by recruiting the active involvement of their consumers in evaluating the social validity of behavioral programs. By "asking them" in meaningful ways, behavioral researchers and their consumers will answer many questions about the "new" social validity and thus determine its validity, social and otherwise.

REFERENCES

- Baer, D. M., & Schwartz, I. S. (1991). If reliance on epidemiology were to become epidemic, we would need to assess its social validity. *Journal of Applied Behavior Analysis*, *24*, 231-234.
- Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*, 836-841.
- Diaz, C., Starfield, B., Holtzman, N., Mellits, E. D., Hankin, J., Smalky, K., & Benson, P. (1986). Ill health and use of medical care: Community-based assessment of morbidity in children. *Medical Care*, *24*, 848-856.
- Eisen, M., Ware, J. E., Jr., Donald, C. A., & Brook, R. G. (1979). Measuring components of children's health status. *Medical Care*, *17*, 902-921.
- Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis*, *24*, 235-239.
- Finney, J. W., Brophy, C. J., Friman, P. C., Golden, A. S., Richman, G. S., & Ross, A. F. (1990). Promoting parent-provider interaction during young children's health-supervision visits. *Journal of Applied Behavior Analysis*, *23*, 207-213.
- Finney, J. W., Riley, A. W., & Cataldo, M. F. (in press). Psychology in primary health care: Effects of brief targeted therapy on children's medical care utilization. *Journal of Pediatric Psychology*.
- Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis*, *24*, 205-213.
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*, 336-352.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12-19.
- Kazdin, A. E. (1977). Assessing the clinical or implied importance of behavior change through social validation. *Behavior Modification*, *1*, 427-452.
- McKinlay, J. B. (1981). From "promising report" to "standard procedure": Seven stages in the career of a

- medical innovation. *Milbank Memorial Fund Quarterly/Health and Society*, **59**, 374-411.
- Schwartz, I. S. (1991). The study of consumer behavior and social validity: An essential partnership for applied behavior analysis. *Journal of Applied Behavior Analysis*, **24**, 241-244.
- Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, **24**, 189-204.
- Winett, R. A., Moore, J. F., & Anderson, E. S. (1991). Extending the concept of social validity: Behavior analysis for disease prevention and health promotion. *Journal of Applied Behavior Analysis*, **24**, 215-230.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, **11**, 203-214.
- Wolf, M. M., & Ramp, K. K. (1991). *Consumer feedback and the Teaching-Family Model: On the auld art of keeping dragons away*. Unpublished manuscript, Department of Human Development and Family Life, University of Kansas, Lawrence.

Received March 3, 1991

Initial editorial decision March 4, 1991

Revision received March 8, 1991

Final acceptance March 8, 1991

Action Editor, E. Scott Geller